AD/A-002 244

# A NOTE ON OPTIMAL STOPPING FOR SUCCESS RUNS

Sheldon M. Ross

California University

Prepared for:

Office of Naval Research
Army Research Office

November 1974

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

| REPORT DOCUMENTATION PAGE | | READ INSTRUCTIONS BEFORE COMPLETING FORM |
|---|---|---|
| 1. REPORT NUMBER<br>ORC 74-33 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER<br>AD/A-002 244 |
| 4. TITLE (and Subtitle)<br>A NOTE ON OPTIMAL STOPPING FOR SUCCESS RUNS | | 5. TYPE OF REPORT & PERIOD COVERED<br>Research Report |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>Sheldon M. Ross | | 8. CONTRACT OR GRANT NUMBER(s)<br>N00014-69-A-0200-1036 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Operations Research Center<br>University of California<br>Berkeley, California 94720 | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS<br>NR 042 238 |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research<br>Department of the Navy<br>Arlington, Virginia 22217 | | 12. REPORT DATE<br>November 1974 |
| | | 13. NUMBER OF PAGES<br>10 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>Unclassified |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Also supported by the U. S. Army Research Office-Durham under Grant DAHC04 74 G 0226.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

One-Stage Look Ahead
Optimal Stopping
Modified Problem



20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

i

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

# A NOTE ON OPTIMAL STOPPING FOR SUCCESS RUNS†

Operations Research Center Research Report No. 74-33

Sheldon M. Ross

November 1974

U. S. Army Research Office - Durham

DAHCO4 74 G 0226

Department of Industrial Engineering
and Operations Research
University of California, Berkeley



†Partially supported by the Office of Naval Research under Contract N00014-69-A-0200-1036 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.



10

THE FINDINGS IN THIS REPORT ARE NOT TO BE CONSTRUED AS AN OFFICIAL DEPARTMENT OF THE ARMY POSITION, UNLESS SO DESIGNATED BY OTHER AUTHORIZED DOCUMENTS.

ABSTRACT

The following model is considered by Starr (1972): At most n tosses of a coin, having a constant probability p of coming up heads, are made. After each toss we have the option of either stopping and receiving an amount equal to the length of the terminal run of heads (that is, if we were on a streak of k heads in the last k tosses, then we could stop and receive k) , or of paying an amount c and tossing the coin again. When n tosses have already been made, we must stop.

The purpose of this note is to point out that with a simple modification the above problem fits the framework in which a one-stage look ahead policy is optimal. This yields not only an easy solution to the problem but also provides much insight. For instance, the reason for the additivity of the optimal continuation boundary, which is commented on by Starr (1972) on Page 1890, now becomes clear. Also the problem may be generalized so that the terminal payoff is a more general function of the terminal run of heads, which may even also depend on the number of tosses made.

/-

# A NOTE ON OPTIMAL STOPPING FOR SUCCESS RUNS

by

Sheldon M. Ross

## 1. INTRODUCTION

The following model is considered by Starr (1972): At most $n$ tosses of a coin, having a constant probability $p$ of coming up heads, are made. After each toss we have the option of either stopping and receiving an amount equal to the length of the terminal run of heads (that is, if we were on a streak of $k$ heads in the last $k$ tosses, then we could stop and receive $k$), or of paying an amount $c$ and tossing the coin again. When $n$ tosses have already been made, we must stop.

The purpose of this note is to point out that with a simple modification the above problem fits the framework in which a one-stage look ahead policy is optimal. This yields not only an easy solution to the problem but also provides much insight. For instance, the reason for the additivity of the optimal continuation boundary, which is commented on by Starr (1972) on Page 1890, now becomes clear. Also the problem may be generalized so that the terminal payoff is a more general function of the terminal run of heads, which may even also depend on the number of tosses made.

## 2. THE OPTIMAL POLICY

Consider the above problem with the exception that the return when we stop after a terminal run of $r$ heads is $f(r)$, where $f(r)$ is such that

$$f(r) - pf(r + 1)$$

is nondecreasing in $r$.

Define $V_n$ to be the value to the decision-maker if he is allowed to make at most $n$ tosses before stopping and when he employs an optimal strategy, and note that $V_n$ is nondecreasing in $n$. Say that the process is in state $(r,j)$ if we are on a run of $r$ heads and we are allowed at most $j$ more coin tosses.

Now let us consider a modified problem which is such that when we are in any state of the form $(0,j)$, $j \geq 0$, we are forced to stop and we receive a terminal reward $V_j$. (That is, whenever a tail occurs we must stop but we are paid as if we acted optimally from this point on.) In this modified problem if we stop when in state $(r,j)$ we receive $f(r)$, while if we continue for exactly one more toss and then stop then our expected return is $pf(r + 1) + (1 - p)V_{j-1} - c$. Hence, the one-stage look ahead policy (see Derman and Sacks (1960), Chow and Robbins (1961) or [3], pp. 137-138) is to stop at state $(r,j)$ either if $r = 0$ or $r \neq 0$ and

$$f(r) \geq pf(r + 1) + (1 - p)V_{j-1} - c .$$

As the set of stopping states just defined is closed in the sense that once entered is never left, it follows that the one-stage look ahead policy is optimal for this modified problem (in the terminology of [1] we are in the monotone case).

As an optimal policy for the modified problem clearly cannot lead to non-optimal actions in states $(r,j)$, $r > 0$, for the original problem, it remains only to determine the optimal actions at states of the form $(0,j)$. To do so we fix $j$ and consider a modified problem, allowing at most $j$ flips and such that we are forced to stop whenever we enter a state $(0,i)$ when $i < j$ and we receive a terminal reward $V_i$. The one-stage look ahead policy for this problem (which, as before, is easily shown to be optimal) calls for stopping at $(0,j)$ if

$$f(0) \geq pf(1) + (1 - p)V_{j-1} - c .$$

Combining this with our previous results shows that for the original problem it is optimal to stop at $(r,j)$ if and only if

$$f(r) - pf(r + 1) \geq (1 - p)V_{j-1} - c .$$

If $f(0) = 0$, then the above states that we should stop if and only if the present payoff $(f(r))$ is at least the expected payoff if we make exactly one more toss $(pf(r + 1) - c)$ plus $1 - p$ times the value of a new game which allows at most $j - 1$ tosses $((1 - p)V_{j-1})$.

## 3. A GENERALIZATION

The problem can be generalized to allow the terminal reward to depend not only on the length of the terminal run of heads but also on the number of tosses taken. That is, assuming that we can initially make at most $n$ tosses then the return if we stop when in state $(r,j)$ would be some function $f(r, n-j)$, $j \leq n$. If the function $f(r,i)$ satisfies

$$(1) \quad \begin{aligned} & f(r,i) \leq f(r, i+1) \\ & f(r+1, i+1) - pf(r+2, i+2) \geq f(r,i) - pf(r+1, i+1) \end{aligned}$$

then it can be shown by the same method as used in Section 2 that it is optimal to stop at $(r,j)$ if and only if

$$f(r, n-j) \geq pf(r+1, n-j+1) + (1-p)V_n(j-1) - c$$

when $V_n(j)$ is the conditional expected return under an optimal policy from time $n-j$ onward given that the head run is of length zero after $n-j$ tosses. An example of a terminal reward satisfying (1) is $f(r,i) = r/i$, $r \leq i$. In words, the terminal reward would equal the terminal head run divided by the number of tosses made.

REFERENCES

[1] Chow, Y. S. and H. Robbins, "A Martingale Systems Theorem and Applications," PROCEEDINGS OF THE FOURTH BERKELEY SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, University of California Press, Vol. 1, pp. 93-104, (1961).

[2] Derman, C. and J. Sacks, "Replacement of Periodically Inspected Equipment (An Optimal Stopping Rule)," Naval Research Logistics Quarterly, Vol. 7, pp. 597-607, (1960).

[3] Ross, S., APPLIED PROBABILITY MODELS WITH OPTIMIZATION APPLICATIONS, Holden-Day, (1970).

[4] Starr, N., "How to Win a War If You Must: Optimal Stopping Based on Success Runs," Annals of Mathematical Statistics, Vol. 6, No. 43, pp. 1884-1893, (1972).